Unsupervised User Intent Modeling by Feature-Enriched Matrix Factorization

Yun-Nung (Vivian) Chen, Ming Sun, Alexander I. Rudnicky, and Anatole Gershman

Carnegie Mellon

Summary

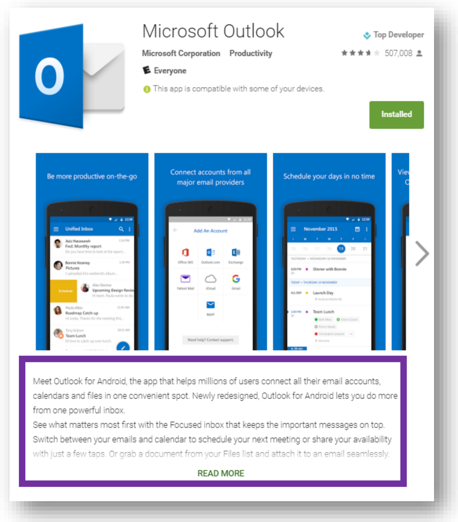
- Challenge of typical SDS: ***Predefined Ontology & Hidden Semantics***
  1) Predefined domain ontology is required to support corresponding functionality
     - Structured knowledge resources are available (e.g. Freebase, Wikipedia, FrameNet) and may provide semantic information
  2) Hidden semantics may contain important semantics
     - Implicit information helps infer feature relations
- Approach: ***Feature-Enriched MF-SLU***
  - Enrich semantics with the structured knowledge for improving intent prediction
  - A single matrix integrating different-level knowledge for reasoning and prediction simultaneously
- Result
  - Feature-enriched MF-SLU benefits from hidden information and rich features, and outperforms the baseline that uses a language-modeling retrieval model.

1. Feature-Enriched MF-SLU: Spoken Language Understanding by Matrix Factorization

- **Data**: speech data collected from users, with intents from 13 frequently accessed domains in Google Play (WER = 19.8%)

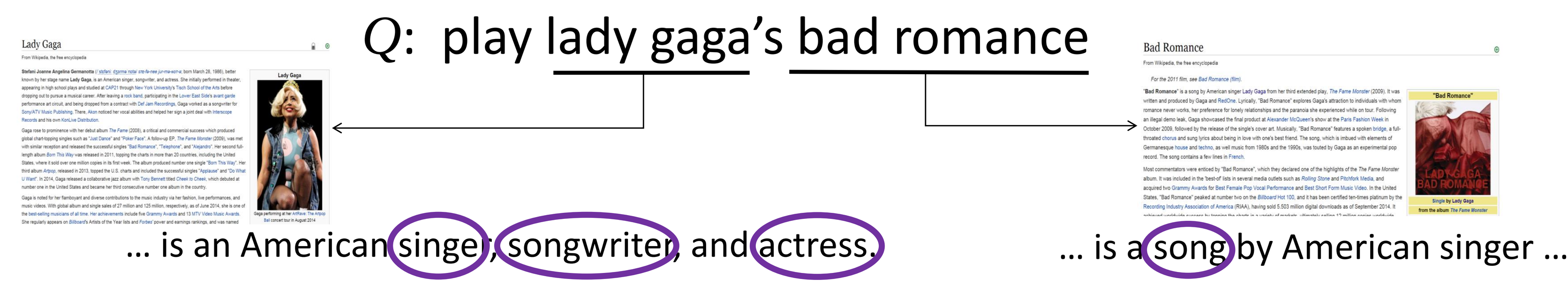
Lexical Matrix

- **Main idea: use manually authored app description as it should describe the app's functionality**



Enriched Semantics Matrix

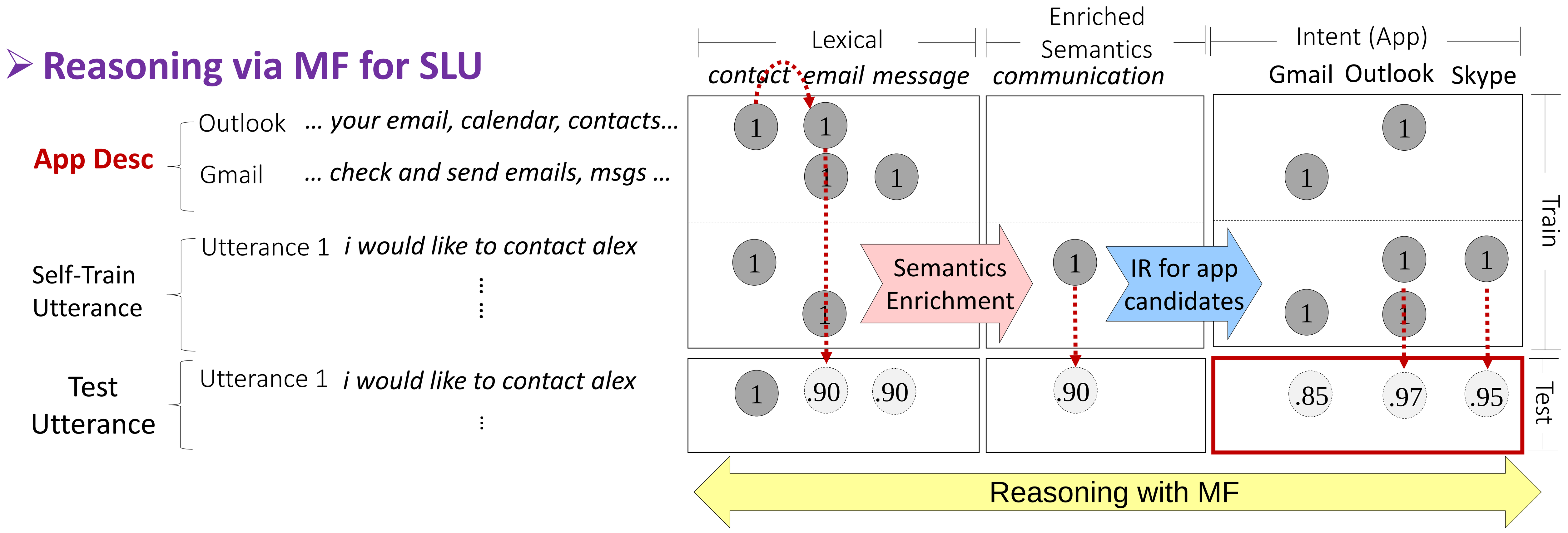
- **Main idea: slot types and word embeddings help infer semantics for expanding domain knowledge**
  - Entity Type from Structured Knowledge (e.g. Wikipedia/Freebase)



Intent Matrix

- **Main idea: retrieve the apps that are most likely to support users' requests, for self-training**

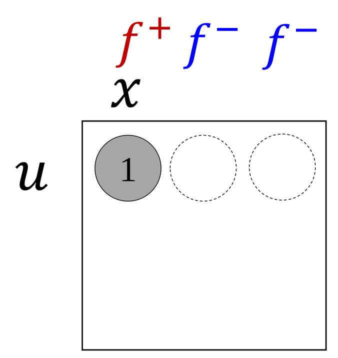
Reasoning via MF for SLU



Chen and Rudnicky, "Dynamically Supporting Unexplored Domains in Conversational Interactions by Enriching Semantics with Neural Word Embeddings," in Proc. of SLT, 2014.

2. Model Learning by Matrix Factorization

- Modeling Implicit Feedback:



$$f^+ = \langle u, x^+ \rangle \quad f^- = \langle u, x^- \rangle \Rightarrow p(f^+) > p(f^-)$$

$$p(M_{u,x} = 1 \mid \theta_{u,x}) = \sigma(\theta_{u,x}) = \frac{1}{1+\exp(-\theta_{u,x})}$$

model parameters

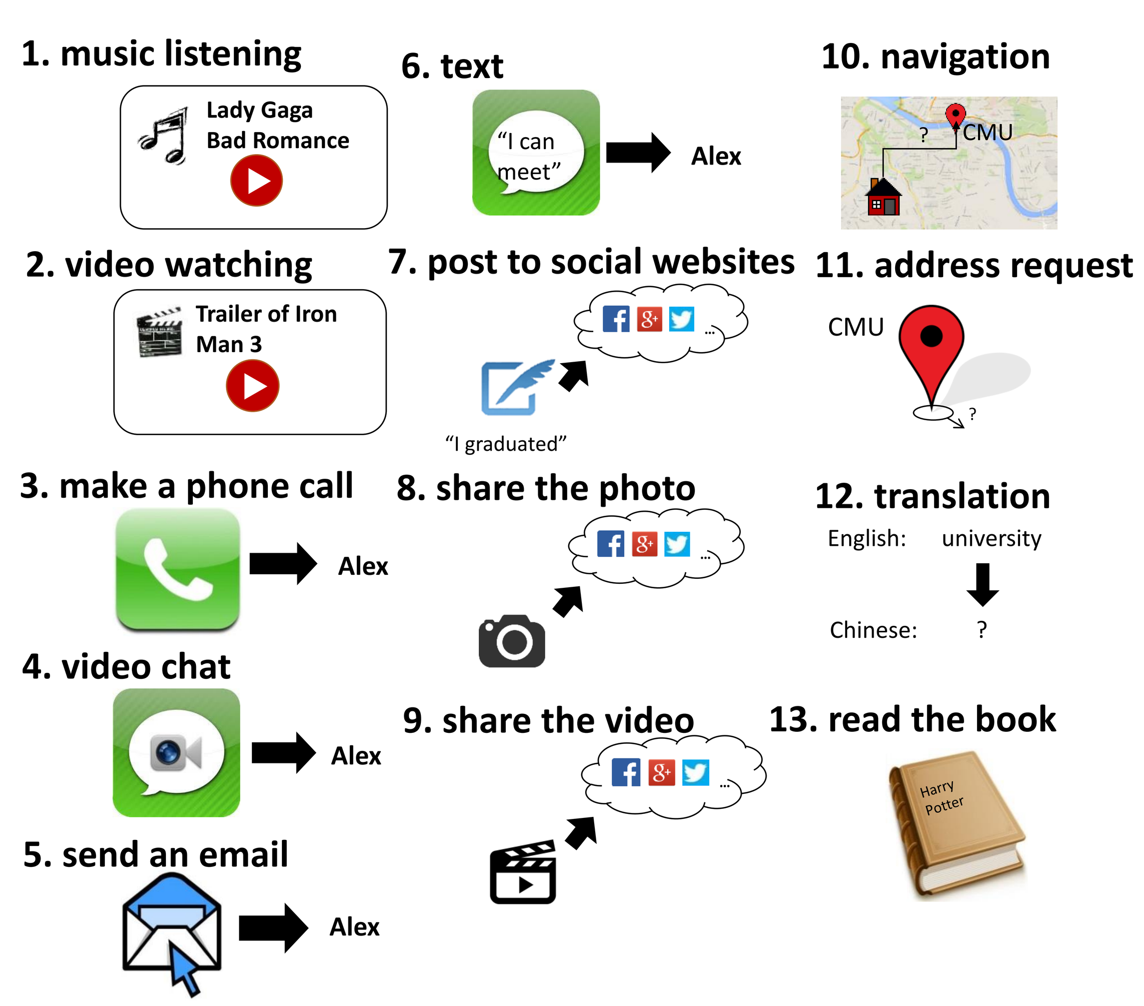
- Objective:

$$\sum_{f^+ \in \mathcal{O}} \sum_{f^- \notin \mathcal{O}} \ln \sigma(\theta_{f^+} - \theta_{f^-})$$

➢ MF learns a set of well-ranked intents per utterance.

3. Experiments

- Dataset: single-turn request with intents below
- Evaluation Metrics
  - Mean Average Precision (MAP)
  - Precision at K (P@K)



- MAP for Intent Modeling

Feature Matrix (MAP)	ASR		Transcripts	
	LM	MF-SLU	LM	MF-SLU
Word Observation	25.1	29.2 (+16.2%)	26.1	30.4 (+16.4%)
+ Embedding-Enriched Semantics	**32.0**	**34.2 (+6.8%)**	**33.3**	33.3 (-0.2%)
+ Type-Embedding-Enriched Semantics	31.5	32.2 (+2.1%)	32.9	**34.0 (+3.4%)**

➢ Enriched semantics significantly improve the performance for intent modeling

- P@10 for Intent Modeling

Feature Matrix (P@10)	ASR		Transcripts	
	LM	MF-SLU	LM	MF-SLU
Word Observation	28.6	29.5 (+3.4%)	29.2	30.1 (+2.8%)
+ Embedding-Enriched Semantics	31.2	**32.5 (+4.3%)**	32.0	33.0 (+3.4%)
+ Type-Embedding-Enriched Semantics	**31.3**	30.6 (-2.3%)	**32.5**	**34.7 (+6.8%)**

➢ Type information inferred from ASR results may not be accurate enough; noisy enriched information could be degrading performance.
➢ When there are no recognition errors, accurate type information benefits performance.

➢ The *feature-enriched MF-SLU* can benefit from both

1) ***hidden information*** *modeled by MF*

2) ***enriched semantics*** *including structured knowledge from different modalities*

to improve Intent prediction.

Conclusion

- We propose an MF approach to learn user intents based on rich feature patterns from multiple modalities, including app descriptions, automatically acquired knowledge and user utterances.
- In a smart-phone intelligent assistant setting (e.g. requesting an app), the feature-enriched MF-SLU can handle users' open domain intents by returning relevant apps that provide desired functionality either locally available or by suggesting installation of suitable apps in an unsupervised way.
- The framework can flexibly extend to incorporate different-level features for improving a system's ability to assist users pursuing personalized multi-app activities.
- The effectiveness of the feature-enriched MF-SLU model can be shown for different domains, indicating good generality and provides a promising direction for future work.